PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [5] :
C12N 15/53, C11B 1/04
A01H 1/04, C12Q 1/68

A1

(11) International Publication Number: WO 91/18985

(43) International Publication Date: 12 December 1991 (12.12.91)

(21) International Application Number: PCT/US91/03288

(22) International Filing Date: 16 May 1991 (16.05.91)

(30) Priority data:
529,049 25 May 1990 (25.05.90) US

(60) Parent Application or Grant
(63) Related by Continuation
US 529,049 (CIP)
Filed on 25 May 1990 (25.05.90)

(71) Applicant *(for all designated States except US)*: E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : HITZ, William, D. [US/US]; 2501 West 18th Street, Wilmington, DE 19806 (US). YADAV, Narendra, S. [US/US]; 127 Jade Drive, Wilmington, DE 19810 (US).

(74) Agents: MORRISSEY, Bruce, W. et al.; E.I. du Pont de Nemours and Company, Legal/Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), BR, CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: NUCLEOTIDE SEQUENCE OF SOYBEAN STEAROYL-ACP DESATURASE GENE

(57) Abstract

The preparation and use of nucleic acid fragments encoding soybean seed stearoyl-ACP desaturase enzyme or its precursor to modify plant oil composition are described. Chimeric genes incorporating such nucleic acid fragments and suitable regulatory sequences may be utilized to transform plants to control the levels of saturated and unsaturated fatty acids.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

TITLE

NUCLEOTIDE SEQUENCE OF SOYBEAN STEAROYL-ACP DESATURASE GENE

BACKGROUND OF THE INVENTION

Soybean oil accounts for about 70% of the 14 billion pounds of edible oil consumed in the United States and is a major edible oil worldwide. It is used in baking, frying, salad dressing, margarine, and a multitude of processed foods. In 1987/88 60 million acres of soybean were planted in the U.S. Soybean is the lowest-cost producer of vegetable oil, which is a by-product of soybean meal. Soybean is agronomically well-adapted to many parts of the U.S. Machinery and facilities for harvesting, storing, and crushing are widely available across the U.S. Soybean products are also a major element of foreign trade since 30 million metric tons of soybeans, 25 million metric tons of soybean meal, and 1 billion pounds of soybean oil were exported in 1987/88. Nevertheless, increased foreign competition has lead to recent declines in soybean acreage and production. The low cost and ready availability of soybean oil provides an excellent opportunity to upgrade this commodity oil into higher value speciality oils to both add value to soybean crop for the U.S. farmer and enhance U.S. trade.

Soybean oil derived from commercial varieties is composed primarily of 11% palmitic (16:0), 4% stearic (18:0), 24% oleic (18:1), 54% linoleic (18:2) and 7% linolenic (18:3) acids. Palmitic and stearic acids are, respectively, 16- and 18-carbon-long saturated fatty acids. Oleic, linoleic and linolenic are 18-carbon-long unsaturated fatty acids containing one, two and three double bonds, respectively. Oleic acid is also referred to as a monounsaturated fatty acid, while linoleic and

linolenic acids are also referred to as polyunsaturated fatty acids. The specific performance and health attributes of edible oils is determined largely by their fatty acid composition.

Soybean oil is high in saturated fatty acids when compared to other sources of vegetable oil and contains a low proportion of oleic acid, relative to the total fatty acid content of the soybean seed. These characteristics do not meet important health needs as defined by the American Heart Association.

More recent research efforts have examined the role that monounsaturated fatty acid plays in reducing the risk of coronary heart disease. In the past, it was believed that monounsaturates, in contrast to saturates and polyunsaturates, had no effect on serum cholesterol and coronary heart disease risk. Several recent human clinical studies suggest that diets high in monounsaturated fat may reduce the "bad" (low-density lipoprotein) cholesterol while maintaining the "good" (high-density lipoprotein) cholesterol. [See Mattson et al. (1985) Journal of Lipid Research 26:194-202, Grundy (1986) New England Journal of Medicine 314:745-748, and Mensink et al.(1987) The Lancet 1:122-125, all collectively herein incorporated by reference.] These results corroborate previous epidemiological studies of people living in Mediterranean countries where a relatively high intake of monounsaturated fat and low consumption of saturated fat correspond with low coronary heart disease mortality. [Keys, A., Seven Countries: A Multivariate Analysis of Death and Coronary Heart Disease, Cambridge: Harvard University Press, 1980, herein incorporated by reference.] The significance of monounsaturated fat in the diet was further confirmed by international researchers from seven countries at the Second Colloquim on

Monounsaturated Fats held February 26, 1987, in Bethesda, MD, and sponsored by the National Heart, Lung and Blood Institutes [Report, Monounsaturates Use Said to Lower Several Major Risk Factors, Food Chemical News, March 2, 1987, p. 44, herein incorporated by reference.]

Soybean oil is also relatively high in polyunsaturated fatty acids -- at levels in far excess of our essential dietary requirement. These fatty acids oxidize readily to give off-flavors and result in reduced performance associated with unprocessed soybean oil. The stability and flavor of soybean oil is improved by hydrogenation, which chemically reduces the double bonds. However, the need for this processing reduces the economic attractiveness of soybean oil.

A soybean oil low in total saturates and polyunsaturates and high in monounsaturate would provide significant health benefits to the United States population, as well as, economic benefit to oil processors. Soybean varieties which produce seeds containing the improved oil will also produce valuable meal as animal feed.

Another type of differentiated soybean oil is an edible fat for confectionary uses. More than 2 billion pounds of cocoa butter, the most expensive edible oil, are produced worldwide. The U.S. imports several hundred million dollars worth of cocoa butter annually. The high and volatile prices and uncertain supply of cocoa butter have encouraged the development of cocoa butter substitutes. The fatty acid composition of cocoa butter is 26% palmitic, 34% stearic, 35% oleic and 3% linoleic acids. About 72% of cocoa butter's triglycerides have the structure in which saturated fatty acids occupy positions 1 and 3 and oleic acid occupies position 2. Cocoa butter's unique fatty acid composition and distribution on the triglyceride

molecule confer on it properties eminently suitable for confectionary end-uses: it is brittle below 27°C and depending on its crystalline state, melts sharply at 25-30°C or 35-36°C. Consequently, it is hard and non-greasy at ordinary temperatures and melts very sharply in the mouth. It is also extremely resistant to rancidity. For these reasons, producing soybean oil with increased levels of stearic acid, especially in soybean lines containing higher-than-normal levels of palmitic acid, and reduced levels of unsaturated fatty acids is expected to produce a cocoa butter substitute in soybean. This will add value to oil and food processors as well as reduce the foreign import of certain tropical oils.

Only recently have serious efforts been made to improve the quality of soybean oil through plant breeding, especially mutagenesis, and a wide range of fatty acid composition has been discovered in experimental lines of soybean (Table 1). These findings (as well as those with other oilcrops) suggest that the fatty acid composition of soybean oil can be significantly modified without affecting the agronomic performance of a soybean plant. However, there is no soybean mutant line with levels of saturates less than those present in commercial canola, the major competitor to soybean oil as a "healthy" oil.

TABLE 1

Range of Fatty Acid Percentages Produced by Soybean Mutants

Fatty Acids	Range of Percentages
Palmitic Acid	6-28
Stearic Acid	3-30
Oleic Acid	17-50
Linoleic Acid	35-60
Linolenic Acid	3-12

There are serious limitations to using mutagenesis to alter fatty acid composition. It is unlikely to discover mutations a) that result in a dominant ("gain-of-function") phenotype, b) in genes that are essential for plant growth, and c) in an enzyme that is not rate-limiting and that is encoded by more than one gene. Even when some of the desired mutations are available in soybean mutant lines their introgression into elite lines by traditional breeding techniques will be slow and expensive, since the desired oil compositions in soybean are most likely to involve several recessive genes.

Recent molecular and cellular biology techniques offer the potential for overcoming some of the limitations of the mutagenesis approach, including the need for extensive breeding. Particularly useful technologies are: a) seed-specific expression of foreign genes in transgenic plants [see Goldberg et al.(1989) Cell 56:149-160], b) use of antisense RNA to inhibit plant target genes in a dominant and tissue-specific manner [see van der Krol et al. (1988) Gene 72:45-50], c) transfer of foreign genes into elite commercial varieties of commercial oilcrops, such as soybean [Chee

et al. (1989) Plant Physiol. 91:1212-1218; Christou et al. (1989) Proc. Natl. Acad. Sci. U.S.A. 86:7500-7504; Hinchee et al. (1988) Bio/Technology 6:915-922; EPO publication 0 301 749 A2], rapeseed [De Block et al. (1989) Plant Physiol. 91:694-701], and sunflower [Everett et al.(1987) Bio/Technology 5:1201-1204], and d) use of genes as restriction fragment length polymorphism (RFLP) markers in a breeding program, which makes introgression of recessive traits into elite lines rapid and less expensive [Tanksley et al. (1989) Bio/Technology 7:257-264]. However, application of each of these technologies requires identification and isolation of commercially-important genes.

Oil biosynthesis in plants has been fairly well-studied [see Harwood (1989) in Critical Reviews in Plant Sciences, Vol. 8(1):1-43]. The biosynthesis of palmitic, stearic and oleic acids occur in the plastids by the interplay of three key enzymes of the "ACP track": palmitoyl-ACP elongase, stearoyl-ACP desaturase and acyl-ACP thioesterase. Stearoyl-ACP desaturase introduces the first double bond on stearoyl-ACP to form oleoyl-ACP. It is pivotal in determining the degree of unsaturation in vegetable oils. Because of its key position in fatty acid biosynthesis it is expected to be an important regulatory step. While the enzyme's natural substrate is stearoyl-ACP, it has been shown that it can, like its counterpart in yeast and mammalian cells, desaturate stearoyl-CoA, albeit poorly [McKeon et al. (1982) J. Biol. Chem. 257:12141-12147]. The fatty acids synthesized in the plastid are exported as acyl-CoA to the cytoplasm. At least three different glycerol acylating enzymes (glycerol-3-P acyltransferase, 1-acyl-glycerol-3-P acyltransferase and diacylglycerol acyltransferase) incorporate the acyl moieties from the cytoplasm into triglycerides during oil biosynthesis.

These acyltransferases show a strong, but not absolute, preference for incorporating saturated fatty acids at positions 1 and 3 and monounsaturated fatty acid at position 2 of the triglyceride. Thus, altering the fatty acid composition of the acyl pool will drive by mass action a corresponding change in the fatty acid composition of the oil. Furthermore, there is experimental evidence that, because of this specificity, given the correct composition of fatty acids, plants can produce cocoa butter substitutes [Bafor et al. (1990) JAOCS 67:217-225].

Based on the above discussion, one approach to altering the levels of stearic and oleic acids in vegetable oils is by altering their levels in the cytoplasmic acyl-CoA pool used for oil biosynthesis. There are two ways of doing this genetically: a) altering the biosynthesis of stearic and oleic acids in the plastid by modulating the levels of stearoyl-ACP desaturase in seeds through either overexpression or antisense inhibition of its gene, and b) converting stearoyl-CoA to oleoyl-CoA in the cytoplasm through the expression of the stearoyl-ACP desaturase in the cytoplasm.

In order to use antisense inhibition of stearoyl-ACP desaturase in the seed, it is essential to isolate the gene(s) or cDNA(s) encoding the target enzyme(s) in the seed, since antisense inhibition requires a high-degree of complementarity between the antisense RNA and the target gene that is expected to be absent in stearoyl-ACP desaturase genes from other species or even in soybean stearoyl-ACP desaturase genes that are not expressed in the seed.

The purification and nucleotide sequences of mammalian microsomal stearoyl-CoA desaturases have been published [Thiede et al. (1986) J. Biol. Chem.

262:13230-13235; Ntambi et al. (1988) J. Biol. Chem. 263:17291-17300; Kaestner et al. (1989) J. Biol. Chem. 264:14755-14761]. However, the plant enzyme differs from them in being soluble, in utilizing a different electron donor, and in its substrate-specificities. The purification and the nucleotide sequences for animal enzymes do not teach how to purify the plant enzyme or isolate a plant gene. The purification of stearoyl-ACP desaturase was reported from safflower seeds [McKeon et al. (1982) J. Biol. Chem. 257:12141-12147]. However, this purification scheme was not useful for soybean, either because the desaturases are different or because of the presence of other proteins such as the soybean seed storage proteins in seed extracts.

The rat liver stearoyl-CoA desaturase protein has been expressed in E. coli [Strittmatter et al. (1988) J. Biol. Chem. 263:2532-2535] but, as mentioned above, its substrate specificity and electron donors are quite distinct from that of the plant.

SUMMARY OF THE INVENTION

A means to control the levels of saturated and unsaturated fatty acids in edible plant oils has been discovered. Utilizing the soybean seed stearoyl-ACP desaturase cDNA for either the precursor or enzyme, chimeric genes are created and may be utilized to transform various plants to modify the fatty acid composition of the oil produced. Specifically, one aspect of the present invention is a nucleic acid fragment comprising a nucleotide sequence encoding the soybean seed stearoyl-ACP desaturase cDNA corresponding to the nucleotides 1 to 2243 in SEQ ID NO:1, or any nucleic acid fragment substantially homologous therewith. Preferred are those nucleic acid fragments encoding the soybean seed stearoyl-ACP desaturase

precursor or the mature soybean seed stearoyl-ACP desaturase enzyme.

Another aspect of this invention involves a chimeric gene capable of transforming a soybean plant cell comprising a nucleic acid fragment encoding the soybean seed stearoyl-ACP desaturase cDNA operably linked to suitable regulatory sequences producing antisense inhibition of soybean seed stearoyl-ACP desaturase in the seed. Preferred are those chimeric genes which incorporate nucleic acid fragments encoding the soybean seed stearoyl-ACP desaturase precursor or the mature soybean seed stearoyl-ACP desaturase enzyme.

Yet another embodiment of the invention involves a method of producing seed oil containing modified levels of saturated and unsaturated fatty acids comprising: (a) transforming a plant cell with a chimeric gene described above, (b) growing sexually mature plants from said transformed plant cells, (c) screening progeny seeds from said sexually mature plants for the desired levels of stearic acid, and (d) crushing said progeny seed to obtain said oil containing modified levels of stearic acid. Preferred plant cells and oils are derived from soybean, rapeseed, sunflower, cotton, cocoa, peanut, safflower, and corn. Preferred methods of transforming such plant cells would include the use of Ti and Ri plasmids of *Agrobacterium*, electroporation, and high-velocity ballistic bombardment.

DETAILED DESCRIPTION OF THE INVENTION

The present invention describes a nucleic acid fragment that encodes soybean seed stearoyl-ACP desaturase. This enzyme catalyzes the introduction of a double bond between carbon atoms 9 and 10 of stearoyl-ACP to form oleoyl-ACP. It can also convert stearoyl-CoA into oleoyl-CoA, albeit with reduced efficiency. Transfer of the nucleic acid fragment of the invention,

or a part thereof that encodes a functional enzyme, with suitable regulatory sequences into a living cell will result in the production or over-production of stearoyl-ACP desaturase, which in the presence of an appropriate electron donor, such as ferredoxin, may result in an increased level of unsaturation in cellular lipids, including oil, in tissues when the enzyme is absent or rate-limiting.

Occasionally, reintroduction of a gene or a part thereof into a plant results in the inhibition of both the reintroduced and the endogenous gene, Jorgenson (December, 1990) Trends in Biotechnology 340-344. Therefore, reintroduction of the nucleic acid fragment of the invention is also expected to, in some cases, result in inhibition of the expression of endogenous seed stearoyl-ACP desaturase and would then result in increased level of saturation in seed oil.

Transfer of the nucleic acid fragment of the invention into a soybean plant with suitable regulatory sequences that transcribe the antisense RNA complementary to the mRNA, or its precursor, for seed stearoyl-ACP desaturase may result in the inhibition of the expression of the endogenous stearoyl-ACP desaturase gene and, consequently, in reduced desaturation in the seed oil.

The nucleic acid fragment of the invention can also be used as a restriction fragment length polymorphism marker in soybean genetic studies and breeding programs.

In the context of this disclosure, a number of terms shall be utilized. As used herein, the term "nucleic acid" refers to a large molecule which can be single stranded or double stranded, composed of monomers (nucleotides) containing a sugar, phosphate and either a purine or pyrimidine. A "nucleic acid fragment" is a fraction of a given nucleic acid molecule. In higher

plants, deoxyribonucleic acid (DNA) is the genetic material while ribonucleic acid (RNA) is involved in the transfer of the information in DNA into proteins. A "genome" is the entire body of genetic material contained in each cell of an organism. The term "nucleotide sequence" refers to a polymer of DNA or RNA which can be single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases capable of incorporation into DNA or RNA polymers. As used herein, the term "homologous to" refers to the complementarity between the nucleotide sequence of two nucleic acid molecules or between the amino acid sequences of two protein molecules. Estimates of such homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those skilled in the art [as described in Hames and Higgins, Eds. (1985) Nucleic Acid Hybridisation, IRL Press, Oxford, U.K.]; or by the comparison of sequence similarity between two nucleic acids or proteins. As used herein, "substantially homologous" refers to nucleic acid molecules which require less stringent conditions of hybridization than those for homologous sequences, and coding DNA sequence which may involve base changes that do not cause a change in the encoded amino acid, or which involve base changes which may alter an amino acid, but not affect the functional properties of the protein encoded by the DNA sequence.

Thus, the nucleic acid fragments described herein include molecules which comprise possible variations of the nucleotide bases derived from deletion, rearrangement, random or controlled mutagenesis of the nucleic acid fragment, and even occasional nucleotide sequencing errors so long as the DNA sequences are substantially homologous.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding) and following (3' non-coding) the coding region. "Stearoyl-ACP desaturase gene" refers to a nucleic acid fragment that expresses a protein with stearoyl-ACP desaturase activity. "Native" gene refers to the gene as found in nature with its own regulatory sequences. "Chimeric" gene refers to a gene that comprises heterogeneous regulatory and coding sequences. "Endogenous" gene refers to the native gene normally found in its natural location in the genome. A "foreign" gene refers to a gene not normally found in the host organism but that is introduced by gene transfer.

"Coding sequence" refers to a DNA sequence that codes for a specific protein and excludes the non-coding sequences. It may constitute an "uninterrupted coding sequence", i.e., lacking an intron, such as in a cDNA or it may include one or more introns bounded by appropriate splice junctions. An "intron" is a sequence of RNA which is transcribed in the primary transcript but which is removed through cleavage and re-ligation of the RNA within the cell to create the mature mRNA that can be translated into a protein.

"Translation initiation codon" and "translation termination codon" refer to a unit of three adjacent nucleotides in a coding sequence that specifies initiation and chain termination, respectively, of protein synthesis (mRNA translation). "Open reading frame" refers to the amino acid sequence encoded between translation initiation and termination codons of a coding sequence.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect

complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA" (mRNA) refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to an RNA transcript that includes the mRNA. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene by interfering with the processing, transport and/or translation of its primary transcript or mRNA. The complementarity of an antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. In addition, as used herein, antisense RNA may contain regions of ribozyme sequences that may increase the efficacy of antisense RNA to block gene expression. "Ribozyme" refers to a catalytic RNA and includes sequence-specific endoribonucleases.

As used herein, "suitable regulatory sequences" refer to nucleotide sequences located upstream (5'), within, and/or downstream (3') to a coding sequence, which control the transcription and/or expression of the coding sequences, potentially in conjunction with the protein biosynthetic apparatus of the cell. In artificial DNA constructs, regulatory sequences can also control the transcription and stability of antisense RNA.

"Promoter" refers to a DNA sequence in a gene, usually upstream (5') to its coding sequence, which controls the expression of the coding sequence by

providing the recognition for RNA polymerase and other factors required for proper transcription. In artificial DNA constructs promoters can also be used to transcribe antisense RNA. Promoters may also contain DNA sequences that are involved in the binding of protein factors which control the effectiveness of transcription initiation in response to physiological or developmental conditions. It may also contain enhancer elements. An "enhancer" is a DNA sequence which can stimulate promoter activity. It may be an innate element of the promoter or a heterologous element inserted to enhance the level and/or tissue-specificity of a promoter. "Constitutive promoters" refers to those that direct gene expression in all tissues and at all times. "Tissue-specific" or "development-specific" promoters as referred to herein are those that direct gene expression almost exclusively in specific tissues, such as leaves or seeds, or at specific development stages in a tissue, such as in early or late embryogenesis, respectively. "Inducible promoters" refers to those that direct gene expression in response to an external stimulus, such as light, heat-shock and chemical.

The term "expression", as used herein, is intended to mean the production of a functional end-product. In the case of expression or overexpression of the stearoyl-ACP desaturase genes it involves transcription of the gene and translation of the mRNA into precursor or mature stearoyl-ACP desaturase proteins. In the case of antisense inhibition it refers to the production of antisense RNA transcripts capable of preventing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms.

The "3' non-coding sequences" refers to that the DNA sequence portion of a gene that contains a polyadenylation signal and any other regulatory signal capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor.

"Mature" protein refers to a functional desaturase enzyme without its transit peptide. "Precursor" protein refers to the mature protein with a native or foreign transit peptide. "Transit" peptide refers to the amino terminal extension of a polypeptide, which is translated in conjunction with the polypeptide forming a precursor peptide and which is required for its uptake by plastids of a cell.

"Transformation" herein refers to the transfer of a foreign gene into the genome of a host organism and its genetically stable inheritance. "Restriction fragment length polymorphism" refers to different sized restriction fragment lengths due to altered nucleotide sequences in or around variant forms of genes, and may be abbreviated as "RFLP". "Fertile" refers to plants that are able to propagate sexually.

Purification of Soybean Seed Stearoyl-ACP Desaturase

Stearoyl-ACP desaturase protein was purified to near-homogeneity from the soluble fraction of extracts made from developing soybean seeds following its chromatography on Blue Sepharose, anion-exchange, alkyl-ACP sepharose, and chromatofocussing on Mono P (Pharmacia). Because of the lability of the enzyme during purification, the nearly homogenous preparation is purified only ca. a few hundred-fold; the basis of this lability is not understood. Chromatofocussing resolved the enzyme into two peaks of activity: the peak

that eluted earlier, with an apparent pI of ca. 6, had a higher specific-activity than the peak eluting later, with an apparent pI of ca. 5.7. The native molecular weight of the purified enzyme was estimated by gel filtration to be ca. 65 kD. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of the purified desaturase preparation showed it to be a polypeptide of ca. 38 kD, which suggests that the native enzyme is a dimer. A smaller polypeptide is occasionally observed in varying amounts resulting in a doublet in some preparations. This appears to be due to a proteolytic breakdown of the larger one, since the level of the smaller one increases during storage. However, it cannot be ruled out that the enzyme could also be a heterodimer or that there are different-sized isozymes.

A highly purified desaturase preparation was resolved on SDS-PAGE, electrophoretically transfered onto Immobilon®-P membrane (Millipore), and stained with Coomassie blue. The ca. 38 kD protein on the Immobilon®-P was cut out and used to make polyclonal antibody in mice.

A C_4 reverse-phase HPLC column was used to further purify the enzyme that eluted earlier in chromatofocussing. The major protein peak was homogeneous for the ca. 38 kD polypeptide. It was used for determining the N-terminal sequence: Arg-Ser-Gly-Ser-Lys-Glu-Val-Glu-Asn-Ile-Lys-Lys-Pro-Phe-Thr-Pro (SEQ ID NO:3).

Cloning of Soybean Seed Stearoyl-ACP Desaturase cDNA

Based on the N-terminal sequence of the purified desaturase protein, a set of eight degenerate 35 nucleotide-long oligonucleotides was designed for use as a hybridization probe. The design took into account the codon usage in selected soybean seed genes and used five

deoxyinosines at selected positions of ambiguity. The probe, following radiolabeling, was used to screen a cDNA expression library made in Lambda ZAP vector from poly A^+ RNA from 20-day old developing soybean seeds. Six positively-hybridizing plaques were subjected to plaque purification. Sequences of the pBluescript (Stratagene) vector, including the cDNA inserts, from each of six purified phages were excised in the presence of a helper phage and the resultant phagemids used to infect E. coli cells resulting in a double-stranded plasmids, pDS1 to pDS6.

The cDNA insert in plasmid pDS1 is flanked at one end (the 5' end of the coding sequence) by the unique Eco RI site and at its other end by the unique Hind III site. Both Eco RI and the Hind III sites are from the vector, pBluescript. The nucleotide sequence of the cDNA insert in pDS1 revealed an open reading frame for 402 amino acids that included the mature protein's N-terminal sequence 43 amino acid residues from the N-terminus of the open reading frame (SEQ ID NO:1). At least part of this "presequence" is the transit peptide required for precursor import into the chloroplast. Although there are four methionines in this presequence that are in-frame with the mature protein sequence, the most likely N-terminal residue is methionine at position -32 (with the N-terminal Arg of mature protein being referred to as +1) since: a) the N-terminal methionine in the transit peptide sequences for all known chloroplast precursor proteins, with only one exception, is followed by alanine, and b) the methionine at position -5 is too close to the N-terminus of the mature protein to be the initiating codon for the transit peptide (the smallest transit sequence found thus far is 31 amino acids long). Thus, it can be deduced that the desaturase precursor protein consists of a 32-amino acid

long transit peptide and a 359-amino acid long mature protein. Based on fusion-protein studies in which the C-terminus of foreign proteins is fused either to the desaturase precursor at position -10 (Ser) or to the mature desaturase protein at position +10 (Ile), the N-terminus of a functional stearoyl-ACP desaturase enzyme can range at least ± 10 amino acids from Arg at position +1 (SEQ ID NO:1).

The restriction maps of all six plasmids, though not identical, showed a common 0.7 kb Bgl II fragment found within the coding region of the precursor for stearoyl-ACP desaturase in pDS1. This strongly suggests that all six clones encode for the stearoyl-ACP desaturase. The partial restriction maps of plasmids pDS1, pDS5 and pDS6 appear to be the identical. The inserts in pDS2 and pDS3, which differ in their physical maps from each other as well as from that of pDS1, were partially sequenced. Their partial nucleotide sequences, including 262 nucleotides from the 3' non-coding region, were identical to that in pDS1.

Of the several cDNA clones isolated from the soybean cDNA library using pDS1 as hybridization probe, five were sequenced in the 3' non-coding sequence and their sequences compared to that of SEQ ID NO:1. The results are summarized below:

Clone #	Sequence correspondence to SEQ ID NO:1	Percent Identity
1	1291-1552	100
2	1291-1394	100
3	1285-1552	100
4	1285-1552	100
5	1298-1505	91

Thus, while the claimed sequence (SEQ ID NO:1) most likely represents the predominantly-expressed stearoyl-

ACP desaturase gene in soybean seed, at least one other stearoyl-ACP desaturase gene that is 91% homologous at the nucleotide level to the claimed sequence. The partial sequence of clone #5 is shown in SEQ ID NO:2.

As expected, comparision of the deduced amino-acid sequences for soybean stearoyl-ACP desaturase and the rat microsomal stearoyl-CoA desaturases did not reveal any significant homology.

In vitro recombinant DNA techniques were used to make two fusion proteins:

a) a recombinant plasmid pGEXB that encodes a ca. 66 kD fusion protein consisting of a 28 kD glutathione-S-transferase (GST) protein fused at its C-terminus to the ca. 38 kD desaturase precursor protein at amino acid residue -10 from the N-terminus of the mature enzyme (Arg, +1) (SEQ ID NO:1). Extracts of *E. coli* cells harboring pGEXB, grown under conditions that induce the synthesis of the fusion protein, show stearoyl-ACP desaturase activity and expression of a ca. 66 kD fusion protein that cross-reacts with antibody made against soybean stearoyl-ACP desaturase and that binds to glutathione-agarose affinity column. The affinity column can be used to purify the fusion protein to near-homogeneity in a single step. The desaturase moiety can be cleaved off in the presence of thrombin and separated from the GST by re-chromatography on the glutathione-agarose column; and

b) a recombinant plasmid, pNS2, that encodes a ca. 42 kD fusion protein consisting of 4 kD of the N-terminus of β-galactosidase fused at its C-terminus to the amino acid residue at position +10 (Ile) from the N-terminus of the mature desaturase protein (Arg, +1) (SEQ ID NO:1). Extract of *E. coli* cells harboring pNS2 express a ca. 42 kD protein that cross-reacts with

antibody made against soybean stearoyl-ACP desaturase and show stearoyl-ACP desaturase activity.

E. coli (pGEXB) can be used to purify the stearoyl-ACP desaturase for use in structure-function studies on the enzyme, in immobilized cells or in extracellular desaturations [see Ratledge et al. (1984) Eds., Biotechnology for the Oils and Fats Industry, American Oil Chemists' Society]. E coli (pNS2) can be used to express the desaturase enzyme in vivo. However, for in vivo function it may be necessary to introduce an electron donor, such as ferredoxin and NADPH:ferredoxin reductase. The ferredoxin gene has been cloned from a higher plant [Smeekens et al. (1985) Nucleic Acids Res. 13:3179-3194] and human ferredoxin has been expressed in E. coli [Coghlan et al. (1989) Proc. Natl. Acad. Sci. USA, 86:835-839]. Alternatively, one skilled in the art can express the mature protein in microorganisms using other expression vectors described in the art [Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press; Milman (1987) Meth. Enzymol. 153:482-491; Duffaud et al. (1987) Meth. Enzymol. 153:492-507; Weinstock (1987) Meth. Enzymol. 154:156-163; E.P.O. Publication 0 295 959 A2).

The fragment of the instant invention may be used, if desired, to isolate substantially homologous stearoyl-ACP desaturase cDNAs and genes, including those from plant species other than soybean. Isolation of homologous genes is well-known in the art. Southern blot analysis reveals that the soybean cDNA for the enzyme hybridizes to several, different-sized DNA fragments in the genomic DNA of tomato, rapeseed (Brassica napus), soybean, corn (a monocotyledenous plant) and Arabidopsis (which has a very simple genome). The Southern blot of corn DNA reveals that the soybean cDNA can also hybridize non-specifically, which may make

the isolation of the corn gene more difficult. Although we do not know how many different genes or "pseudogenes" (non-functional genes) are present in any plant, it is expected to be more than one, since stearoyl-ACP desaturase is an important enzyme. Moreover, plants that are amphidiploid (that is, derived from two progenitor species), such as soybean, rapeseed (B. napus), and tobacco will have genes from both progenitor species.

The nucleic acid fragment of the instant invention encoding soybean seed stearoyl-ACP desaturase cDNA, or a coding sequence derived from other cDNAs or genes for the enzyme, with suitable regulatory sequences, can be used to overexpress the enzyme in transgenic soybean as well as other transgenic species. Such a recombinant DNA construct may include either the native stearoyl-ACP desaturase gene or a chimeric gene. One skilled in the art can isolate the coding sequences from the fragment of the invention by using and/or creating sites for restriction endonucleases, as described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. Of particular utility are sites for Nco I (5'-CCATGG-3') and Sph I (5'-GCATGC-3') that allow precise removal of coding sequences starting with the initiating codon ATG. The fragment of invention has a Nco I recognition sequence at nucleotide positions 1601-1606 (SEQ ID NO:1) that is 357 bp after the termination codon for the coding sequence. For isolating the coding sequence of stearoyl-ACP desaturase precursor from the fragment of the invention, an Nco I site can be engineered by substituting nucleotide A at position 69 with C. This will allow isolation of the 1533 bp Nco I fragment containing the precursor coding sequence. The expression of the mature enzyme in the cytoplasm is

expected to desaturate stearoyl-CoA to oleoyl-CoA. For this it may be necessary to also express the mature ferredoxin in the cytoplasm, the gene for which has been cloned from plants [Smeekens et al. (1985) Nucleic Acids Res. 13:3179-3194]. For isolating the coding sequence for the mature protein, a restriction site can be engineered near nucleotide position 164. For example, substituting nucleotide G with nucleotide C at position 149 or position 154 would result in the creation of Nco I site or Sph I site, respectively. This will allow isolation of a 1453 bp Nco I fragment or a 1448 bp Sph I-Nco I fragment, each containing the mature protein sequence. Based on fusion protein studies, the N-terminus of the mature stearoyl-ACP desaturase enzyme is not critical for enzyme activity.

Antisense RNA has been used to inhibit plant target genes in a dominant and tissue-specific manner [see van der Krol et al. (1988) Gene 72:45-50; Ecker et al. (1986) Proc. Natl. Acad. Sci. USA 83:5372-5376; van der Krol et al. (1988) Nature 336:866-869; Smith et al. (1988) Nature 334:724-726; Sheehy et al. (1988) Proc. Natl. Acad. Sci. USA 85:8805-8809; Rothstein et al. (1987) Proc. Natl. Acad. Sci. USA 84:8439-8443; Cornelissen et al. (1988) Nucl. Acids Res. 17:833-843; Cornelissen (1989) Nucl. Acid Res. 17:7203-7209; Robert et al. (1989) Plant Mol. Biol. 13:399-409].

The use of antisense inhibition of the seed enzyme would require isolation of the coding sequence for genes that are expressed in the target tissue of the target plant. Thus, it will be more useful to use the fragment of the invention to screen seed-specific cDNA libraries, rather than genomic libraries or cDNA libraries from other tissues, from the appropriate plant for such sequences. Moreover, since there may be more than one gene encoding seed stearoyl-ACP desaturase, it may be

useful to isolate the coding sequences from the other genes from the appropriate crop. The genes that are most highly expressed are the best targets for antisense inhibition. The level of transcription of different genes can be studied by known techniques, such as run-off transcription.

For expressing antisense RNA in soybean seed from the fragment of the invention, the entire fragment of the invention (that is, the entire cDNA for soybean stearoyl-ACP desaturase from the unique Eco RI to Hind III sites in plasmid pDS1) may be used. There is evidence that the 3' non-coding sequences can play an important role in antisense inhibition [Ch'ng et al.(1989) Proc. Natl. Acad. Sci. USA 86:10006-10010]. There have also been examples of using the entire cDNA sequence for antisense inhibition [Sheehy et al. (1988) Proc. Natl. Acad. Sci. USA 89:8439-8443]. The Hind III and Eco RI sites can be modified to facilitate insertion of the sequences into suitable regulatory sequences in order to express the antisense RNA.

A preferred host soybean plant for the antisense RNA inhibition of stearoyl-ACP desaturase for producing a cocoa butter substitute in soybean seed oil is a soybean plant containing higher-than-normal levels of palmitic acid, such as A19 double mutant, which is being commercialized by Iowa State University Research Foundation, Inc. (315 Beardshear, Ames, Iowa 50011).

A preferred class of heterologous hosts for the expression of the coding sequence of stearoyl-ACP desaturase precursor or the antisense RNA are eukaryotic hosts, particularly the cells of higher plants. Particularly preferred among the higher plants are the oilcrops, such as soybean (Glycine max), rapeseed (Brassica napus, B. campestris), sunflower (Helianthus annus), cotton (Gossypium hirsutum), corn (Zea mays),

cocoa (Theobroma cacao), and peanut (Arachis hypogaea). Expression in plants will use regulatory sequences functional in such plants.

The expression of foreign genes in plants is well-established [De Blaere et al.(1987) Meth. Enzymol. 153:277-291]. The origin of promoter chosen to drive the expression of the coding sequence or the antisense RNA is not critical as long as it has sufficient transcriptional activity to accomplish the invention by increasing or decreasing, respectively, the level of translatable mRNA for stearoyl-ACP desaturase in the desired host tissue. Preferred promoters include strong plant promoters (such as the constitutive promoters derived from Cauliflower Mosaic Virus that direct the expression of the 19S and 35S viral transcripts [Odell et al.(1985) Nature 313:810-812; Hull et al. (1987) Virology 86:482-493]), small subunit of ribulose 1,5-bisphosphate carboxylase [Morelli et al.(1985) Nature 315:200; Broglie et al.(1984) Science 224:838; Hererra-Estrella et al.(1984) Nature 310:115; Coruzzi et al.(1984) EMBO J. 3:1671; Faciotti et al.(1985) Bio/Technology 3:241], maize zein protein [Matzke et al.(1984) EMBO J. 3:1525], and chlorophyll a/b binding protein [Lampa et al.(1986) Nature 316:750-752].

Depending upon the application, it may be desirable to select inducible promoters and/or tissue- or development-specific promoters. Such examples include the light-inducible promoters of the small subunit of ribulose 1,5-bisphosphate carboxylase genes (if the expression is desired in tissues with photosynthetic function).

Particularly preferred tissue-specific promoters are those that allow seed-specific expression. This may be especially useful, since seeds are the primary source of vegetable oils and also since seed-specific

expression will avoid any potential deleterious effect in non-seed tissues. Examples of seed-specific promoters include but are not limited to the promoters of seed storage proteins, which can represent up to 90% of total seed protein in many plants. The seed storage proteins are strictly regulated, being expressed almost exclusively in seeds in a highly tissue-specific and stage-specific manner [Higgins et al. (1984) Ann. Rev. Plant Physiol. 35:191-221; Goldberg et al. (1989) Cell 56:149-160]. Moreover, different seed storage proteins may be expressed at different stages of seed development.

Expression of seed-specific genes has been studied in great detail [see reviews by Goldberg et al. (1989) Cell 56:149-160 and Higgins et al. (1984) Ann. Rev. Plant Physiol. 35:191-221]. There are currently numerous examples for seed-specific expression of seed storage protein genes in transgenic dicotyledonous plants. These include genes from dicotyledonous plants for bean β-phaseolin [Sengupta-Gopalan et al. (1985) Proc. Natl. Acad. Sci. USA 82:3320-3324; Hoffman et al. (1988) Plant Mol. Biol. 11:717-729], bean lectin [Voelker et al. (1987) EMBO J. 6: 3571-3577], soybean lectin [Okamuro et al. (1986) Proc. Natl. Acad. Sci. USA 83: 8240-8244], soybean kunitz trypsin inhibitor [Perez-Grau et al. (1989) Plant Cell 1:095-1109], soybean β-conglycinin [Beachy et al. (1985) EMBO J. 4:3047-3053; Barker et al. (1988) Proc. Natl. Acad. Sci. USA 85:458-462; Chen et al. (1988) EMBO J. 7:297-302; Chen et al. (1989) Dev. Genet. 10:112-122; Naito et al. (1988) Plant Mol. Biol. 11:109-123], pea vicilin [Higgins et al. (1988) Plant Mol. Biol. 11:683-695], pea convicilin [Newbigin et al. (1990) Planta 180:461], pea legumin [Shirsat et al. (1989) Mol. Gen. Genetics 215:326]; rapeseed napin [Radke et al. (1988) Theor. Appl. Genet.

75:685-694] as well as genes from monocotyledonous plants such as for maize 15-kD zein [Hoffman et al. (1987) EMBO J. 6:3213-3221], and barley β-hordein [Marris et al. (1988) Plant Mol. Biol. 10:359-366] and wheat glutenin [Colot et al. (1987) EMBO J. 6:3559-3564]. Moreover, promoters of seed-specific genes operably linked to heterologous coding sequences in chimeric gene constructs also maintain their temporal and spatial expression pattern in transgenic plants. Such examples include Arabidopsis thaliana 2S seed storage protein gene promoter to express enkephalin peptides in Arabidopsis and B. napus seeds [Vandekerckhove et al. (1989) Bio/Technology 7:929-932], bean lectin and bean β-phaseolin promoters to express luciferase [Riggs et al. (1989) Plant Sci. 63:47-57], and wheat glutenin promoters to express chloramphenicol acetyl transferase [Colot et al. (1987) EMBO J. 6:3559-3564].

Of particular use in the expression of the nucleic acid fragment of the invention will be the heterologous promoters from several extensively-characterized soybean seed storage protein genes such as those for the Kunitz trypsin inhibitor [Jofuku et al. (1989) Plant Cell 1:1079-1093; Perez-Grau et al. (1989) Plant Cell 1:1095-1109], glycinin [Nielson et al. (1989) Plant Cell 1:313-328], β-conglycinin [Harada et al. (1989) Plant Cell 1:415-425]. Promoters of genes for α- and β-subunits of soybean β-conglycinin storage protein will be particularly useful in expressing the mRNA or the antisense RNA to stearoyl-ACP desaturase in the cotyledons at mid- to late-stages of seed development [Beachy et al. (1985) EMBO J. 4:3047-3053; Barker et al. (1988) Proc. Natl. Acad. Sci. USA 85:458-462; Chen et al. (1988) EMBO J. 7:297-302; Chen et al. (1989) Dev. Genet. 10:112-122; Naito et al. (1988) Plant Mol. Biol.

11:109-123] in transgenic plants, since: a) there is very little position effect on their expression in transgenic seeds, and b) the two promoters show different temporal regulation: the promoter for the α-subunit gene is expressed a few days before that for the β-subunit gene; this is important for transforming rapeseed where oil biosynthesis begins about a week before seed storage protein synthesis [Murphy et al. (1989) J. Plant Physiol. 135:63-69].

Also of particular use will be promoters of genes expressed during early embryogenesis and oil biosynthesis. The native regulatory sequences, including the native promoter, of the stearoyl-ACP desaturase gene expressing the nucleic acid fragment of the invention can be used following its isolation by those skilled in the art. Heterologous promoters from other genes involved in seed oil biosynthesis, such as those for _B. napus_ isocitrate lyase and malate synthase [Comai et al. (1989) Plant Cell 1:293-300], _Arabidopsis_ ACP [Post-Beittenmiller et al. (1989) Nucl. Acids Res. 17:1777], _B. napus_ ACP [Safford et al. (1988) Eur. J. Biochem. 174:287-295], _B. campestris_ ACP [Rose et al. (1987) Nucl. Acids Res. 15:7197] may also be used. The partial protein sequences for the relatively-abundant enoyl-ACP reductase and acetyl-CoA carboxylase are published [Slabas et al. (1987) Biochim. Biophys. Acta 877:271-280; Cottingham et al. (1988) Biochim. Biophys. Acta 954: 201-207] and one skilled in the art can use these sequences to isolate the corresponding seed genes with their promoters.

Proper level of expression of stearoyl-ACP mRNA or antisense RNA may require the use of different chimeric genes utilizing different promoters. Such chimeric genes can be transfered into host plants either together

in a single expression vector or sequentially using more than one vector.

It is envisioned that the introduction of enhancers or enhancer-like elements into either the native stearoyl-ACP desaturase promoter or into other promoter constructs will also provide increased levels of primary transcription for antisense RNA or in RNA for stearoyl-ACP desaturase to accomplish the inventions. This would include viral enhancers such as that found in the 35S promoter [Odell et al. (1988) Plant Mol. Biol. 10:263-272], enhancers from the opine genes (Fromm et al. (1989) Plant Cell 1:977-984), or enhancers from any other source that result in increased transcription when placed into a promoter operably linked to the nucleic acid fragment of the invention.

Of particular importance is the DNA sequence element isolated from the gene for the α-subunit of β-conglycinin that can confer 40-fold seed-specific enhancement to a constitutive promoter [Chen et al. (1988) EMBO J. 7:297-302; Chen et al. (1989) Dev. Genet. 10:112-122]. One skilled in the art can readily isolate this element and insert it within the promoter region of any gene in order to obtain seed-specific enhanced expression with the promoter in transgenic plants. Insertion of such an element in any seed-specific gene that is expressed at different times than the β-conglycinin gene will result in expression in transgenic plants for a longer period during seed development.

The invention can also be accomplished by a variety of other methods to obtain the desired end. In one form, the invention is based on modifying plants to produce increased levels of stearoyl-ACP desaturase by virtue of having significantly larger numbers of copies of either the wild-type or a stearoyl-ACP desaturase gene from a different soybean tissue in the plants.

This may result in sufficient increases in stearoyl-ACP desaturase levels to accomplish the invention.

Any 3' non-coding region capable of providing a polyadenylation signal and other regulatory sequences that may be required for the proper expression of the stearoyl-ACP desaturase coding region can be used to accomplish the invention. This would include the native 3' end of the substantially homologous soybean stearoyl-ACP desaturase gene(s), the 3' end from a heterologous stearoyl-ACP desaturase gene, the 3' end from viral genes such as the 3' end of the 35S or the 19S cauliflower mosaic virus transcripts, the 3' end from the opine synthesis genes, the 3' ends of ribulose 1,5-bisphosphate carboxylase or chlorophyll a/b binding protein, or 3' end sequences from any source such that the sequence employed provides the necessary regulatory information within its nucleic acid sequence to result in the proper expression of the promoter/stearoyl-ACP desaturase coding region combination to which it is operably linked. There are numerous examples in the art that teach the usefulness of different 3' non-coding regions.

Various methods of transforming cells of higher plants according to the present invention are available to those skilled in the art (see EPO publications 0 295 959 A2 and 0 318 341 A1). Such methods include those based on transformation vectors based on the Ti and Ri plasmids of Agrobacterium spp. It is particularly preferred to use the binary type of these vectors. Ti-derived vectors transform a wide variety of higher plants, including monocotyledonous and dicotyledonous plants, such as soybean, cotton and rape [Pacciotti et al.(1985) Bio/Technology 3:241; Byrne et al. (1987) Plant Cell, Tissue and Organ Culture 8:3; Sukhapinda et al. (1987) Plant Mol. Biol. 8:209-216;

Lorz et al. (1985) Mol. Gen. Genet. 199:178; Potrykus (1985) Mol. Gen. Genet. 199:183]. Other transformation methods are available to those skilled in the art, such as direct uptake of foreign DNA constructs [see EPO publication 0 295 959 A2], techniques of electroporation [see Fromm et al. (1986) Nature (London) 319:791] or high-velocity ballistic bombardment with metal particles coated with the nucleic acid constructs [see Kline et al. (1987) Nature (London) 327:70]. Once transformed the cells can be regenerated by those skilled in the art.

Of particular relevance are the recently described methods to transform foreign genes into commercially important crops, such as rapeseed [see De Block et al. (1989) Plant Physiol. 91:694-701], sunflower [Everett et al. (1987) Bio/Technology 5:1201], and soybean [McCabe et al. (1988) Bio/Technology 6:923; Hinchee et al. (1988) Bio/Technology 6:915; Chee et al. (1989) Plant Physiol. 91:1212-1218; Christou et al. (1989) Proc. Natl. Acad. Sci USA 86:7500-7504; EPO Publication 0 301 749 A2].

The use of restriction fragment length polymorphism (RFLP) markers in plant breeding has been well-documented in the art [see Tanksley et al. (1989) Bio/Technology 7:257-264]. The nucleic acid fragment of the invention has been mapped to four different loci on a soybean RFLP map [Tingey et al. (1990) J. Cell Biochem., Supplement 14E p. 291, abstract R153]. It can thus be used as a RFLP marker for traits linked to these mapped loci. More preferably these traits will include altered levels of stearic acid. The nucleic acid fragment of the invention can also be used to isolate the stearoyl-ACP desaturase gene from variant (including mutant) soybeans with altered stearic acid levels. Sequencing of these genes will reveal nucleotide

differences from the normal gene that cause the variation. Short oligonucleotides designed around these differences may be used as hybridization probes to follow the variation in stearic and oleic acids. Oligonucleotides based on differences that are linked to the variation may be used as molecular markers in breeding these variant oil traits.

SEQ ID NO:1 represents the nucleotide sequence of a soybean seed stearoyl-ACP desaturase cDNA and the translation reading frame that includes the open reading frame for the soybean seed stearoyl-ACP desaturase. The nucleotide sequence reads from 5' to 3'. Three letter codes for amino acids are used as defined by the Commissioner, 1114 OG 29 (May 15, 1990) incorporated by reference herein. Nucleotide 1 is the first nucleotide of the cDNA insert after the EcoRI cloning site of the vector and nucleotide 2243 is the last nucleotide of the cDNA insert of plasmid pDS1 which encods the soybean seed stearoyl-ACP desaturase. Nucleotides 70 to 72 are the putative translation initiation codon, nucleotides 166 to 168 are the codon for the N-terminal amino acid of the purified enzyme, nucleotides 1243 to 1245 are the termination codon, nucleotides 1 to 69 are the 5' untranslated sequence, and nucleotides 1246 to 2243 are the 3' untranslated nucleotides. SEQ ID NO:2 represents the partial sequence of a soybean seed stearoyl-ACP desaturase cDNA. The first and last nucleotides (1 and 216 on clone 5) are read 5' to 3' and represent the 3' non-coding sequence. SEQ ID NO:3 represents the N-terminal sequence of the purified soybean seed stearoyl-ACP desaturase. SEQ ID NO:4 represents the degenerate coding sequence for amino acids 5 through 16 of SEQ ID NO:3. SEQ ID NO:5 represents a complimentary mixture of degenerate oligonucleotides to SEQ ID NO:4.

The present invention is further defined in the following EXAMPLES, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these EXAMPLES, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these EXAMPLES, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLE 1
ISOLATION OF cDNA FOR SOYBEAN SEED STEAROYL-ACP DESATURASE

PREPARATION OF [9,10-^{3}H]-STEAROYL-ACP

Purification of Acyl Carrier Protein (ACP) from E. coli

To frozen E. coli cell paste, (0.5 kg of 1/2 log phase growth of E. coli B grown on minimal media and obtained from Grain Processing Corp, Muscatine, IA) was added 50 mL of a solution 1 M in Tris, 1 M in glycine, and 0.25 M in EDTA. Ten mL of 1 M $MgCl_2$ was added and the suspension was thawed in a water bath at 50°C. As the suspension approached 37°C it was transferred to a 37°C bath, made to 10 mM in 2-mercaptoethanol and 20 mg of DNAse and 50 mg of lysozyme were added. The suspension was stirred for 2 h, then sheared by three 20 second bursts in a Waring Blendor. The volume was adjusted to 1 L and the mixture was centrifuged at 24,000xg for 30 min. The resultant supernatant was centrifuged at 90,000xg for 2 h. The resultant high-speed pellet was saved for extraction of acyl-ACP

synthase (see below) and the supernatant was adjusted to pH 6.1 by the addition of acetic acid. The extract was then made to 50% in 2-propanol by the slow addition of cold 2-propanol to the stirred solution at 0°C. The resulting precipitate was allowed to settle for 2 h and then removed by centrifugation at 16,000xg. The resultant supernatant was adjusted to pH 6.8 with KOH and applied at 2 mL/min to a 4.4 x 12 cm column of DEAE-Sephacel which had been equilibrated in 10 mM MES, pH 6.8. The column was washed with 10 mM MES, pH 6.8 and eluted with 1 L of a gradient of LiCl from 0 to 1.7 M in the same buffer. Twenty mL fractions were collected and the location of eluted ACP was determined by applying 10 µL of every second fraction to a lane of a native polyacrylamide (20% acrylamide) gel electrophoresis (PAGE). Fractions eluting at about 0.7 M LiCl contained nearly pure ACP and were combined, dialyzed overnight against water and then lyophilized.

Purification of Acyl-ACP Synthase

Membrane pellets resulting from the high-speed centrifugation described above were homogenized in 380 mL of 50 mM Tris-Cl, pH 8.0, and 0.5 M in NaCl and then centrifuged at 80,000xg for 90 min. The resultant supernatant was discarded and the pellets resuspended in 50 mM Tris-Cl, pH 8.0, to a protein concentration of 12 mg/mL. The membrane suspension was made to 2% in Triton X-100 and 10 mM in $MgCl_2$, and stirred at 0°C for 20 min before centrifugation at 80,000xg for 90 min. The protein in the resultant supernatant was diluted to 5 mg/mL with 2% Triton X-100 in 50 mM Tris-Cl, pH 8.0 and, then, made to 5 mM ATP by the addition of solid ATP (disodium salt) along with an equimolar amount of $NaHCO_3$. The solution was warmed in a 55°C bath until the internal temperature reached 53°C and was then

maintained at between 53°C and 55°C for 5 min. After 5 min the solution was rapidly cooled on ice and centrifuged at 15,000xg for 15 min. The supernatant from the heat treatment step was loaded directly onto a column of 7 mL Blue Sepharose 4B which had been equilibrated in 50 mM Tris-Cl, pH 8.0, and 2% Triton X-100. The column was washed with 5 volumes of the loading buffer, then 5 volumes of 0.6 M NaCl in the same buffer and the activity was eluted with 0.5 M KSCN in the same buffer. Active fractions were assayed for the synthesis of acyl-ACP, as described below, combined, and bound to 3 mL settled-volume of hydroxylapatite equilibrated in 50 mM Tris-Cl, pH 8.0, 2% Triton X-100. The hydroxylapatite was collected by centrifugation, washed twice with 20 mL of 50 mM Tris-Cl, pH 8.0, 2% Triton X-100. The activity was eluted with two 5 mL washes of 0.5 M potassium phosphate, pH 7.5, 2% Triton X-100. The first wash contained 66% of the activity and it was concentrated with a 30 kD membrane filtration concentrator (Amicon) to 1.5 mL.

<u>Synthesis of [9,10-^{3}H]-Stearoyl-ACP</u>

A solution of stearic acid in methanol (1 mM, 34.8 µL) was mixed with a solution of [9,10-^{3}H]stearate (Amersham) containing 31.6 µCi of ^{3}H and dried in a glass vial. The ACP preparation described above (1.15 mL, 32 nmoles) was added along with 0.1 mL of 0.1 M ATP, 0.05 mL of 80 mM DTT, 0.1 mL of 8 M LiCl, and 0.2 mL of 13% Triton X-100 in 0.5 M Tris-Cl, pH 8.0, with 0.1 M $MgCl_2$. The reaction was mixed thoroughly and 0.3 mL of the acyl-ACP synthase preparation was added. After 1 h at 37°C, a 10 µL aliquot was taken and dried on a small filter paper disc. The disc was washed extensively with chloroform:methanol:acetic acid (8:2:1, v:v:v) and radioactivity retained on the disc was taken as a

measure of stearoyl-ACP. At 1 h about 67% of the ACP had been consumed and the reaction did not proceed further in the next 2 h. The reaction mix was diluted 1 to 4 with 20 mM Tris-Cl, pH 8.0, and applied to a 1 mL DEAE-Sephacel column equilibrated in the same buffer. The column was washed in sequence with 5 mL of 20 mM Tris-Cl, pH 8.0, 5 mL of 80% 2-propanol in 20 mM Tris-Cl, pH 8.0, and eluted with 0.5 M LiCl in 20 mM Tris-Cl, pH 8.0. The column eluate was passed directly onto a 3 mL column of octyl-sepharose CL-4B which was washed with 10 mL of 20 mM potassium phosphate, pH 6.8, and then eluted with 35% 2-propanol in 2 mM potassium phosphate, pH 6.8. The eluted volume (5.8 mL) contained 14.27 µCi of ^3H (49% yield based on ACP). The eluted product was lyophilized and redissolved at a concentration of 24 µM [^3H]stearoyl-ACP at 0.9 mCi/µmole.

PREPARATION OF ALKYL-ACP AFFINITY COLUMN

Synthesis of N-hexadecyliodoacetamide

1-Hexadecylamine (3.67 mmole) was dissolved in 14.8 mL of CH_2Cl_2, cooled to 4°C, and 2.83 mmoles of iodoacetic anhydride in 11.3 mL of CH_2Cl_2 was added dropwise to the stirred solution. The solution was warmed to room temperature and held for 2 h. The reaction mixture was diluted to about 50 mL with CH_2Cl_2 and washed 3 times (25 mL) with saturated sodium bicarbonate solution and then 2 times with water. The volume of the solution was reduced to about 5 mL under vacuum and passed through 25 mL of silica in diethyl ether. The eluate was reduced to an off-white powder under vacuum. This yielded 820 mg (2.03 mmoles) of the N-hexadecyliodoacetamide (71.8% yield). The 300 MHz ^1H NMR spectra of the product was consistent with the expected structure.

Synthesis of N-Hexadecylacetamido-S-ACP

E. coli ACP prepared as above (10 mg in 2 mL of 50 mM Tris-Cl, pH 7.6) was treated at 37°C with 50 mM DTT for 2 h. The solution was made to 10% TCA, held at 0°C for 20 min and centrifuged to pellet. The resultant pellet was washed (2 x 2 mL) with 0.1 M citrate, pH 4.2 and redissolved in 3 mL of 50 mM potassium phosphate buffer. The pH of the ACP solution was adjusted to 7.5 with 1 M KOH and 3 mL of *N*-hexadecyliodoacetamide (3 mM in 2-propanol) was added. A slight precipitate of the *N*-hexadecyliodoacetamide was redissolved by warming the reaction mix to 45°C. The mixture was held at 45°C for 6 h. SDS-PAGE on 20% acrylamide PAGE gel showed approximately 80% conversion to an ACP species of intermediate mobility between the starting, reduced ACP and authentic palmitoyl-ACP. Excess *N*-hexadecyliodoacetamide was removed from the reaction mix by 4 extractions (3 mL) with CH_2Cl_2 with gentle mixing to avoid precipitation of the protein at the interface.

Coupling of N-Hexadecylacetamido-S-ACP to CNBr-activated Sepharose 4B

Cyanogen bromide-activated Sepharose 4B (Pharmacia, 2 g) was suspended in 1 mM HCl and extensively washed by filtration and resuspension in 1 mM HCl and finally one wash in 0.1 M $NaHCO_3$, pH 8.3. The N-hexadecylacetamido-S-ACP prepared above was diluted with an equal volume of 0.2 M $NaHCO_3$, pH 8.3. The filtered cyanogen bromide-activated Sepharose 4B (about 5 mL) was added to the N-hexadecylacetamido-S-ACP solution, the mixture was made to a volume of 10 mL with the 0.1 M $NaHCO_3$, pH 8.3, and mixed by tumbling at room temperature for 6 h. Protein remaining in solution (Bradford assay) indicated approximately 85% binding. The gel suspension was

collected by centrifugation, washed once with the 0.1 M $NaHCO_3$, pH 8.3, and resuspended in 0.1 M ethanolamine adjusted to pH 8.5 with HCl. The suspension was allowed to stand at 4°C overnight and then washed by centrifugation and re-suspension in 12 mL of 0.1 M acetate, pH 4.0, 0.5 M in NaCl and then 0.1 M $NaHCO_3$, pH 8.3, 0.5 M in NaCl. The alkyl-ACP Sepharose 4B was packed into a 1 x 5.5 cm column and washed extensively with 20 mM bis-tris propane-Cl (BTP-Cl), pH 7.2, before use.

STEAROYL-ACP DESATURASE ASSAY

Stearoyl-ACP desaturase was assayed as described by McKeon et al. [(1982) J. Biol. Chem. 257:12141-12147] except for using [9,10-^{3}H]-stearoyl-ACP. Use of the tritiated substrate allowed assaying the enzyme activity by release of tritium as water, although the assay based on the tritium release underestimates desaturation by a factor of approximately 4 relative to that observed using ^{14}C-stearoyl-ACP by the method of McKeon et al. [(1982) J. Biol. Chem 257:12141-12147], apparently because not all tritium is at carbons 9 and 10. Nevertheless, this modification makes the enzyme assay more sensitive, faster and more reliable. The reaction mix consisted of enzyme in 25 µL of 230 µg/mL bovine serum albumin (Sigma), 49 µg/mL catalase (Sigma), 0.75 mM NADPH, 7.25 µM spinach ferredoxin, and 0.35 µM spinach ferredoxin:NADPH$^+$ oxidoreductase, 50 mM Pipes, pH 6.0, and 1 µM [9,10-^{3}H]-stearoyl-ACP (0.9 mCi/µmole). All reagents, except for the Pipes buffer, labeled substrate and enzyme extract, were preincubated in a volume of 7.25 µL at pH 8.0 at room temperature for 10 min before adding 12.75 µL the Pipes buffer and labeled substrate stocks. The desaturase reaction was usually terminated after 5 min by the addition of 400 µL 10%

trichloroacetic acid and 50 μL of 10 mg/mL bovine serum albumin. After 5 min on ice, the protein precipitate was removed by centrifugation at 13,000xg for 5 min. An aliquot of 425 μL was removed from the resultant supernatant and extracted twice with 2 mL of hexane. An aliquot of 375 μL of the aqueous phase following the second hexane extraction was added to 5 mL of ScintiVerse® Bio HP (Fisher) scintillation fluid and used to determine radioactivity released as tritium.

PURIFICATION OF SOYBEAN SEED STEAROYL-ACP DESATURASE

Developing soybean seeds, ca. 20-25 days after flowering, were harvested and stored at -80°C until use. 300 g of the seeds were resuspended in 600 mL of 50 mM BTP-Cl, pH 7.2, and 5 mM dithiothreitol (DTT) in a Waring Blendor. The seeds were allowed to thaw for a few minutes at room temperature to 4°C and all of the purification steps were carried out at 4°C unless otherwise noted. The seeds were homogenized in the blendor three times for 30 s each and the homogenate was centrifuged at 14,000xg for 20 min. The resultant supernatant was centrifuged at 100,000xg for 1 h. The resultant high-speed supernatant was applied, at a flow-rate of 5 mL/min to a 2.5 x 20 cm Blue Sepharose column equilibrated in 10 mM BTP-Cl, pH 7.2, 0.5 mM DTT. Following a wash with 2 column volumes of 10 mM BTP-Cl, pH 7.2, 0.5 mM DTT, the bound proteins were eluted in the same buffer containing 1 M NaCl. The eluting protein peak, which was detected by absorbance at 280 nm, was collected and precipitated with 80% ammonium sulfate. Following collection of the precipitate by centrifugation at 10,000xg for 20 min, its resuspension in 10 mM potassium phosphate, pH 7.2, 0.5 mM DTT, overnight dialysis in the same buffer precipitate, and clarification through a 0.45 micron filter, it was

applied to a 10 mm x 25 cm Wide-pore™ PEI (NH_2) anion-exchange column (Baker) at 3 mL/min thoroughly equilibrated in buffer A (10 mM potassium phosphate, pH 7.2). After washing the column in buffer A until no protein was eluted, the column was subjected to elution by a gradient from buffer A at 0 min to 0.25 M potassium phosphate (pH 7.2) at 66 min at a flow rate of 3 mL/min. Three mL fractions were collected. The desaturase activity eluted in fractions 17-25 (the activity peak eluted at ca. 50 mM potassium phosphate). The pooled fractions were made to 60 mL with buffer A and applied at 1 mL/min to a 1 x 5.5 cm alkyl-ACP column equilibrated in buffer A containing 0.5 mM DTT. After washing the bound protein with the start buffer until no protein was eluted, the bound protein was eluted by a gradient from buffer A containing 0.5 mM DTT at 0 min to 0.5 M potassium phosphate, pH 7.2, 0.5 mM DTT at 60 min and 1 M potassium phosphate, pH 7.2, 0.5 mM DTT. Four mL fractions were collected. Fractions 15-23, which contained the enzyme with the highest specific activity, were pooled and concentrated to 3 mL by a 30 kD Centricon® concentrator (Millipore) and desalted in a small column of G-25 Sephadex® equilibrated with 25 mM bis-Tris-Cl, pH 6.7. The desalted sample was applied at 1 mL/min to a chromatofocussing Mono P HR 5/20 (Pharmacia) column equilibrated with 25 mM bis-Tris-Cl, pH 6.7, washed with a column volume of the same buffer, and eluted with 1:10 dilution of Polybuffer 74 (Pharmacia) made to pH 5.0 with HCl. Desaturase activity eluted in two peaks: one in fraction 30 corresponding to a pI of ca. 6.0 and the other in fraction 35, corresponding to a pI of ca. 5.7. The protein in the two peaks were essentially composed of ca. 38 kD polypeptide. The first peak had a higher enzyme specific activity and was used for further

characterization as well as for further purification on reverse-phase chromatography.

Mono P fractions containing the first peak of enzyme activity were pooled and applied to a C_4 reverse-phase HPLC column (Vydac) equilibrated with buffer A (5% acetonitrile, 0.1% trifluoroacetic acid) and eluted at 0.1 mL/min with a gradient of 25% buffer B (100% acetonitrile, 0.1% trifluoroacetic acid) and 75% buffer A at 10 min to 50% buffer B and 50% buffer A at 72.5 min. A single major peak eluted at 41.5% buffer B that also ran as a ca. 38 kD protein based on SDS-PAGE. The protein in the peak fraction was used to determine the N-terminal amino acid sequence on a Applied Biosystems 470A Gas Phase Sequencer. The PTH amino acids were analysed on Applied Biosystems 120 PTH Amino Acid Analyzer.

The N-terminal sequence of the ca. 38 kD polypeptide was determined through 16 residues and is shown in SEQ ID NO:3.

CLONING OF SOYBEAN SEED STEAROYL-ACP DESATURASE cDNA

Based on the N-terminal amino acid sequence of the purified soybean seed stearoyl-ACP desaturase (SEQ ID NO:3), amino acids 5 through 16, which are represented by the degenerate coding sequence, SEQ ID NO:4, was chosen to design the complementary mixture of degenerate oligonucleotides (SEQ ID NO:5).

The design took into account the codon bias in representative soybean seed genes encoding Bowman-Birk protease inhibitor [Hammond et al. (1984) J. Biol. Chem. 259:9883-9890], glycinin subunit A-2B-1a (Utsumi et al. (1987) Agric. Biol. Chem. 51:3267-3273], lectin (le-1)[Vodkin et al. (1983) Cell 34:1023-1031], and lipoxygenase-1 [Shibata et al. (1987) J. Biol. Chem.

262:10080-10085]. Five deoxyinosines were used at selected positions of ambiguity.

A cDNA library was made as follows: Soybean embryos (ca. 50 mg fresh weight each) were removed from the pods and frozen in liquid nitrogen. The frozen embryos were ground to a fine powder in the presence of liquid nitrogen and then extracted by Polytron homogenization and fractionated to enrich for total RNA by the method of Chirgwin et al. [Biochemistry (1979) 18:5294-5299]. The nucleic acid fraction was enriched for poly A^+ RNA by passing total RNA through an oligo-dT cellulose column and eluting the poly A^+ RNA by salt as described by Goodman et al. [(1979) Meth. Enzymol. 68:75-90]. cDNA was synthesized from the purified poly A^+ RNA using cDNA Synthesis System (Bethesda Research Laboratory) and the manufacturer's instructions. The resultant double-stranded DNA was methylated by DNA methylase (Promega) prior to filling-in its ends with T4 DNA polymerase (Bethesda Research Laboratory) and blunt-end ligating to phosphorylated Eco RI linkers using T4 DNA ligase (Pharmacia). The double-stranded DNA was digested with Eco RI enzyme, separated from excess linkers by passing through a gel filtration column (Sepharose CL-4B), and ligated to Lambda ZAP vector (Stratagene) as per manufacturer's instructions. Ligated DNA was packaged into phage using Gigapack packaging extract (Stratagene) according to manufacturer's instructions. The resultant cDNA library was amplified as per Stratagene's instructions and stored at -80°C.

Following the instructions in Lambda ZAP Cloning Kit Manual (Stratagene), the cDNA phage library was used to infect E. coli BB4 cells and plated to yield ca. 80,000 plaques per petri plate (150 mm diameter). Duplicate lifts of the plates were made onto

nitrocellulose filters (Schleicher & Schuell). Duplicate lifts from five plates were prehybridized in 25 mL of Hybridization buffer consisting of 6X SSC (0.9 M NaCl, 0.09 M sodium citrate, pH 7.0), 5X Denhardt's [0.5 g Ficoll (Type 400, Pharmacia), 0.5 g polyvinylpyrrolidone, 0.5 g bovine serum albumin (Fraction V; Sigma)], 1 mM EDTA, 1% SDS, and 100 ug/mL denatured salmon sperm DNA (Sigma Chemical Co.) at 45°C for 10 h. Ten pmol of the hybridization probe (see above) were end-labeled in a 52.5 uL reaction mixture containing 50 mM Tris-Cl, pH 7.5, 10 mM $MgCl_2$, 0.1 mM spermidine-HCl (pH 7.0), 1 mM EDTA (pH 7.0), 5 mM DDT, 200 uCi (66.7 pmoles) of gamma-labeled $AT^{32}P$ (New England Nuclear) and 25 units of T4 polynucleotide kinase (New England Biolabs). After incubation at 37°C for 45 min, the reaction was terminated by heating at 68°C for 10 min. Labeled probe was separated from unincorporated $AT^{32}P$ by passing the reaction through a Quick-Spin™ (G-25 Sephadex®) column (Boehringer Mannheim Biochemicals). The purified labeled probe (1.2×10^7 dpm/pmole) was added to the prehybridized filters, following their transfer to 10 mL of fresh Hybridization buffer. Following incubation of the filters in the presence of the probe for 16 h in a shaker at 48°C, the filters were washed in 200 mL of Wash buffer (6X SSC, 0.1% SDS) five times for 5 min each at room temperature, and then once at 48°C for 5 min. The washed filters were air dried and subjected to autoradiography on Kodak XAR-2 film in the presence of intensifying screens (Lightening Plus, DuPont Cronex®) at -80°C overnight. Six positively-hybridizing plaques were subjected to plaque purification as described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press]. Following the Lambda ZAP Cloning Kit Instruction Manual (Stratagene),

sequences of the pBluescript vector, including the cDNA inserts, from each of six purified phages were excised in the presence of a helper phage and the resultant phagemids were used to infect E. coli XL-1 Blue cells resulting in double-stranded plasmids, pDS1 to pDS6. The restriction maps of all six plasmids, though not identical, showed a common 0.7 kb Bgl II fragment found in the desaturase gene (see below).

DNA from plasmids pDS1-pDS6 were made by the alkaline lysis miniprep procedure described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. The alkali-denatured double-stranded DNAs were sequenced using Sequenase® T7 DNA polymerase (US Biochemical Corp.) and the manufacturer's instructions. The sequence of the cDNA insert in plasmid pDS1 is shown in SEQ ID NO:1.

EXAMPLE 2
EXPRESSION OF SOYBEAN SEED STEAROYL-ACP DESATURASE IN E. COLI

Construction of Glutathione-S-Transferase: Stearoyl-ACP Desaturase Fusion Protein

Plasmid pDS1 was linearized with Hind III enzyme, its ends filled-in with Klenow fragment (Bethesda Research Laboratory) in the presence of 50 µM each of all four deoxynucleotide triphosphates as per manufacturer's instructions, and extracted with phenol:chloroform (1:1). Phosphorylated Eco RI linkers (New England Biolabs) were ligated to the DNA using T4 DNA ligase (New England Biolabs). Following partial digestion with Bgl II enzyme and complete digestion with excess Eco RI enzyme, the DNA was run on an agarose gel and stained with ethidium bromide. The 2.1 kb DNA

fragment resulting from a partial Bgl II and Eco RI digestion was cut out of the gel, purified using USBioclean™ (US Biochemicals), and ligated to Bam HI and Eco RI cleaved vector pGEX2T [Pharmacia; see Smith et al. (1988) Gene 67:31] using T4 DNA ligase (New England Biolabs). The ligated mixture of DNAs were used to transform E. coli XL-1 blue cells (Stratagene). Transformants were picked as ampicillin-resistant cells and the plasmid DNA from several transformants analyzed by digestion with Bam HI and Eco RI double restriction digest, as described by Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. Plasmid DNA from one transformant, called pGEXB, showed the restriction pattern expected from the correct fusion. The double-stranded plasmid pGEXB was purified and sequenced to confirm the correct fusion by the Sequenase kit (US Biochemical Corp.). The fusion protein consists of a 28 kD glutathione-S-transferase protein fused at its C-terminus to the desaturase precursor protein at Ser at residue -10 from the N-terminus of the mature enzyme (Arg, +1) (SEQ ID NO:1). Thus, it includes ten amino acids from the transit peptide sequence in addition to the mature protein.

Inducible Expression of the Glutathione-S-Transferase-Stearoyl-ACP Desaturase Fusion Protein

Five mL precultures of plasmids pGEXB and pGEX2T, which were grown overnight at 37°C in LB medium [Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press] containing 100 ug/mL ampicillin, were diluted 1:10 in fresh LB medium containing 100 μg/mL ampicillin and continued to grow on a shaker at 37°C for another 90 min before adding isopropylthio-β-D-galactoside and ferric

chloride to final concentrations of 0.3 mM and 50 μM, respectively. After an additional 3 h on a shaker at 37°C, the cultures were harvested by centrifugation at 4,000xg for 10 min at 4°C. The cells were resuspended in one-tenth of the culture volume of freshly-made and ice-cold Extraction buffer (20 mM sodium phosphate, pH 8.0, 150 mM NaCl, 5 mM EDTA and 0.2 mM phenylmethylsulfonyl fluoride) and re-centrifuged as above. The resultant cells were resuspended in 1/50 vol of the culture in Extraction buffer and sonicated for three ten-second bursts. The sonicated extracts were made to 1% in Triton X-100 and centrifuged at 8,000xg for 1 min in Eppendorf Micro Centrifuge (Brinkmann Instruments) to remove the cellular debris. The supernatant was poured into a fresh tube and used for enzyme assays, SDS-PAGE analysis and purification of the fusion protein.

Five μL aliquots of the extracts were assayed for stearoyl-ACP desaturase activity in a 1 min reaction, as described in Example I. The activities [net pmol of stearoyl-ACP desaturated per min per mL of extract; the blank (no desaturase enzyme) activity was 15 pmol/min/ml] are shown below:

Reaction mixture	Net pmol/min/mL
E. coli (pGEX2T)	0
E. coli (pGEXB)	399
E. coli (pGEXB) - NADPH	0
E. coli (pGEXB) - ferredoxin	0
E. coli (pGEXB) - ferredoxin-NADPH reductase	3

These results show that the desaturase enzyme activity is present in the extract of *E. coli* cells containing pGEXB but not in that of cells containing the

control plasmid pGEX2T. Furthermore, this activity was dependent on an exogenous electron donor.

Proteins in extracts of E. coli cells harboring plasmids pGEX2T or pGEXB were resolved by SDS-PAGE, transferred onto Immobilon®-P (Millipore) and cross-reacted with mouse antibody made against purified soybean stearoyl-ACP desaturase, as described by Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. The resultant Western blot showed that pGEXB encodes for ca. 64 kD GST-stearoyl-ACP desaturase fusion polypeptide, although some lower molecular-weight cross-reacting polypeptides can also be observed, which may represent either a degradation or incomplete synthesis of the fusion protein. It is not known whether the GST-desaturase fusion protein is enzymatically active, since the activity observed may be due to the incomplete fusion by the peptides seen here. The fusion polypeptide was not present in extracts of cells harboring the control plasmid (pGEX2T) nor in extracts of cells harboring pGEXB that were not induced by isopropylthio-β-D-galactoside.

Purification of the Glutathione-S-Transferase-Stearoyl-ACP Desaturase Fusion Protein

The GST-desaturase fusion protein was purified in a one step glutathione-agarose affinity chromatography under non-denaturing conditions, following the procedure of Smith et al. [Gene (1988) 67:31]. For this, the bacterial cell extract was mixed with 1 mL glutathione-agarose (sulfur-linkage, Sigma), equilibrated with 20 mM sodium phosphate, pH 8.0, 150 mM NaCl, for 10 min at room temperature. The beads were collected by centrifugation at 1000xg for 1 min, and washed three times with 1 mL of 20 mM sodium phosphate, pH 8.0, 150

mM NaCl (each time the beads were collected by centrifugation as described above). The fusion protein was eluted with 5 mM reduced glutathione (Sigma) in 50 mM Tris-Cl, pH 8.0. The proteins in the eluted fraction were analyzed by SDS-PAGE and consisted of mostly pure ca. 64 kD GST-desaturase polypeptide, 28 kD GST and a trace of ca. 38 kD desaturase polypeptide. The fusion polypeptide was cleaved in the presence of thrombin, as described by Smith et al. [Gene (1988) 67:31].

Construction of β-Galactosidase-Stearoyl-ACP Desaturase Fusion Protein

Plasmid pDS1 DNA was digested with Ssp I and Pvu I enzymes and the digested DNA fragments were resolved by electrophoresis in agarose. The blunt-ended 2.3 kb Ssp I fragment was cut out of the agarose (Pvu I cleaves a contaminating 2.3 kb Ssp I fragment), purified by USBioclean™ (US Biochemical Corp.), and ligated to vector plasmid pBluescript SK (-) (Stratagene) that had previously been filled-in with Klenow fragment (Bethesda Research Laboratory) following linearization with Not I enzyme. The ligated DNAs were transformed into competent *E. coli* XL-1 blue cells. Plasmid DNA from several ampicillin-resistant transformants were analysed by restriction digestion. One plasmid, called pNS2, showed the expected physical map. This plasmid is expected to encode a ca. 42 kD fusion protein consisting of 4 kD N-terminal of β-galactosidase fused at its C-terminus to isoleucine at residue +10 from the N-terminus of the mature desaturase protein (Arg, +1) (SEQ ID NO:1). Thus, it includes all but the first 10 amino acids of the mature protein. Nucleotide sequencing has not been performed on pNS2 to confirm correct fusion.

Five mL of preculture of *E. coli* cells harboring plasmid pNS2 grown overnight in LB medium containing 100

μg/mL ampicillin was added to 50 mL of fresh LB medium with 100 μg/mL ampicillin. After additional 1 h of growth at 37°C in a shaker, isopropylthio-β-D-galactoside and ferric chloride were added to final concentrations of 0.3 mM and 50 μM, respectively. After another 2 h on a shaker at 37°C, the culture was harvested by centrifugation at 4,000xg for 10 min at 4°C. The cells were resuspended in 1 mL of freshly-made and ice-cold TEP buffer (100 mM Tris-Cl, pH 7.5, 10 mM EDTA and 0.1 mM phenylmethylsulfonyl fluoride) and re-centrifuged as above. The cells were resuspended in 1 mL of TEP buffer and sonicated for three ten-second bursts. The sonicates were made to 1% in Triton X-100, allowed to stand in ice for 5 min, and centrifuged at 8,000xg for 1 min in an Eppendorf Micro Centrifuge (Brinkmann Instruments) to remove the cellular debris. The supernatant was poured into a fresh tube and used for enzyme assays and SDS-PAGE analysis.

A 1 μL aliquot of the extract of E. coli cells containing plasmid pNS2 was assayed for stearoyl-ACP desaturase activity in a 5 min reaction, as described above. The extract showed activity of 288 pmol of stearoyl-ACP desaturated per min per ml of the extract [The blank (no desaturase enzyme) activity was 15 pmol/min/mL].

Proteins in the extract of E. coli cells harboring plasmids pNS2 were resolved by SDS-PAGE, transferred onto Immobilon®-P (Millipore) and cross-reacted with mouse antibody made against purified soybean stearoyl-ACP desaturase, as described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. The resultant Western blot showed that pNS2 encodes for ca. 42 kD β-galactosidase-stearoyl-ACP desaturase fusion polypeptide.

EXAMPLE 3

USE OF SOYBEAN SEED STEAROYL-ACP DESATURASE SEQUENCE IN PLASMID pDS1 AS A RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP) MARKER

Plasmid pDS1 was linearized by digestion with restriction enzyme Eco RI in standard conditions as described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press] and labeled with ^{32}P using a Random Priming Kit from Bethesda Research Laboratories under conditions recommended by the manufacturer. The resulting radioactive probe was used to probe a Southern blot [Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press] containing genomic DNA from soybean [*Glycine max* (cultivar Bonus) and *Glycine soja* (PI81762)], digested with one of several restriction enzymes. After hybridization and washes under standard conditions [Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press] autoradiograms were obtained and different patterns of hybridization (polymorphisms) were identified in digests performed with restriction enzymes Pst 1 and Eco RI. The same probe was then used to map the polymorphic pDS1 loci on the soybean genome, essentially as described by Helentjaris et al. [(1986) Theor. Appl. Genet. 72:761-769]. Plasmid pDS1 probe was applied, as described above, to Southern blots of Eco RI or Pst I digested genomic DNAs isolated from 68 F2 progeny plants resulting from a *G. max* Bonus x *G. soja* PI81762 cross. The bands on the autoradiograms were interpreted as resulting from the inheritance of either paternal (Bonus) or maternal (PI81762) pattern, or both (a heterozygote). The resulting data were subjected to genetic analysis using the computer program Mapmaker

[Lander et al., (1987) Genomics 1: 174-181]. In conjunction with previously obtained data for 436 anonymous RFLP markers in soybean [Tingey et al. (1990) J. Cell. Biochem., Supplement 14E p. 291, abstract R153], we were able to position four genetic loci corresponding to the pDS1 probe on the soybean genetic map. This information will be useful in soybean breeding targeted towards developing lines with altered saturate levels, especially for the high stearic acid mutant phenotype, since these recessive traits are most likely be due to loss of seed stearoyl-ACP desaturase enzyme.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Hitz, William D.
Yadav, Narendra S

(ii) TITLE OF THE INVENTION: Nucleotide Sequence of SoybeanStearoyl-ACP Desaturase cDNA

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: E. I. du Pont de Nemours and Company
(B) STREET: 1007 Market Street
(C) CITY: Wilmington
(D) STATE: Delaware
(E) COUNTRY: USA
(F) ZIP: 19898

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: DISKETTE, 3.50 inch, 1.0 MB
(B) COMPUTER: Apple Macintosh
(C) OPERATING SYSTEM:
(D) SOFTWARE:

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 07/529,049
(B) FILING DATE: 25-MAY-1990
(C) CLASSIFICATION:

(vii) ATTORNEY/AGENT INFORMATION;
(A) NAME: Bruce W. Morrissey
(B) REGISTRATION NUMBER: 30,663
(C) REFERENCE/DOCKET NUMBER: BB-1022

(viii) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (302) 892-4927
(B) TELEFAX: (302) 892-7949
(C) TELEX: 835420

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:2243 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: No

(iv) ANTISENSE: No

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Glycine max
(B) STRAIN: Cultivar Wye
(D) DEVELOPMENTAL STAGE: Developing seeds

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: cDNA to mRNA
(B) CLONE: pDS1

(ix) FEATURE:

(A) NAME/KEY:

(i) 5' non-coding sequence

(ii) Putative translation initiation codon

(iii) Putative transit peptide coding sequence

(iv) Mature protein coding sequence

(v) Translation termination codon

(vi) 3' non-coding sequence

(B) LOCATION:

(i) nucleotides 1 through 69

(ii) nucleotides 70 through 72

(iii) nucleotides 70 through 165

(iv) nucleotides 166 through 1242

(v) nucleotides 1243 through 1245

(vi) nucleotides 1246 through 2243

(C) IDENTIFICATION METHOD:

(i) deduced by proximity to ii) below

(ii) similarity of the context of the methionine codon in the open reading frame to translation initiation codons of other plastid transit peptides

(iii) deduced by proximity to ii) above and iv) below

(iv) experimental determination of N-terminal amino acid sequence and subunit size of purified soybean seed stearoyl-ACP desaturase

(v) The translation termination codon ends the open reading frame for a protein of the expected size

(vi) established by proximity to v) above

(D) OTHER INFORMATION:
Extracts of E. coli expressing the mature protein as a fusion protein show stearoyl-ACP desaturase activity and produce a protein that cross-reacts to stearoyl-ACP desaturase antibody

(x) PUBLICATION INFORMATION: Sequence not published.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
          CTTCTACATT ACTCTCTCTT CTCCTAAAAA TTTCTAATGC            40

TTCCATTGCT TCATCTGACT CACTCATCA ATG GCT CTG AGA CTG AAC CCT   90
                                Met Ala Leu Arg Leu Asn Pro
                                -32     -30

ATC CCC ACC CAA ACC TTC TCC CTC CCC CAA ATG CCC AGC CTC AGA  135
Ile Pro Thr Gln Thr Phe Ser Leu Pro Gln Met Pro Ser Leu Arg
-25                 -20                 -15

TCT CCC CGC TTC CGC ATG GCT TCC ACC CTC CGC TCC GGT TCC AAA  180
Ser Pro Arg Phe Arg Met Ala Ser Thr Leu Arg Ser Gly Ser Lys
-10                 -5                  1               5
```

```
GAG GTT GAA AAT ATT AAG AAG CCA TTC ACT CCT CCC AGA GAA GTG   225
Glu Val Glu Asn Ile Lys Lys Pro Phe Thr Pro Pro Arg Glu Val
                 10                  15                  20

CAT GTT CAA GTA ACC CAC TCT ATG CCT CCC CAG AAG ATT GAG ATT   270
His Val Gln Val Thr His Ser Met Pro Pro Gln Lys Ile Glu Ile
                 25                  30                  35

TTC AAA TCT TTG GAG GAT TGG GCT GAC CAG AAC ATC TTG ACT CAT   315
Phe Lys Ser Leu Glu Asp Trp Ala Asp Gln Asn Ile Leu Thr His
                 40                  45                  50

CTT AAA CCT GTA GAA AAA TGT TGG CAA CCA CAG GAT TTT TTA CCC   360
Leu Lys Pro Val Glu Lys Cys Trp Gln Pro Gln Asp Phe Leu Pro
                 55                  60                  65

GAC CCC TCC TCA GAT GGA TTT GAA GAG CAA GTG AAG GAA CTG AGA   405
Asp Pro Ser Ser Asp Gly Phe Glu Glu Gln Val Lys Glu Leu Arg
                 70                  75                  80

GAG AGA GCA AAG GAG ATT CCA GAT GAT TAC TTT GTT GTT CTT GTC   450
Glu Arg Ala Lys Glu Ile Pro Asp Asp Tyr Phe Val Val Leu Val
                 85                  90                  95

GGA GAC ATG ATC ACA GAG GAA GCT CTG CCT ACT TAC CAA ACT ATG   495
Gly Asp Met Ile Thr Glu Glu Ala Leu Pro Thr Tyr Gln Thr Met
                 95                  100                 110

TTA AAT ACT TTG GAT GGA GTT CGT GAT GAA ACA GGT GCC AGC CTT   540
Leu Asn Thr Leu Asp Gly Val Arg Asp Glu Thr Gly Ala Ser Leu
                 115                 120                 125

ACT TCC TGG GCA ATT TGG ACA AGG GCA TGG ACT GCT GAA GAA AAC   585
Thr Ser Trp Ala Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu Asn
                 130                 135                 140

AGA CAC GGT GAT CTT CTT AAC AAA TAT CTG TAC TTG AGT GGA CGA   630
Arg His Gly Asp Leu Leu Asn Lys Tyr Leu Tyr Leu Ser Gly Arg
                 145                 150                 155

GTT GAC ATG AAA CAA ATT GAG AAG ACA ATT CAG TAC CTT ATT GGG   675
Val Asp Met Lys Gln Ile Glu Lys Thr Ile Gln Tyr Leu Ile Gly
                 160                 165                 170

TCT GGG ATG GAT CCT CGG ACC GAG AAC AGC CCC TAC CTT GGT TTC   720
Ser Gly Met Asp Pro Arg Thr Glu Asn Ser Pro Tyr Leu Gly Phe
                 175                 180                 185

ATT TAC ACT TCA TTT CAA GAG AGG GCA ACC TTC ATA TCC CAC GGA   765
Ile Tyr Thr Ser Phe Gln Glu Arg Ala Thr Phe Ile Ser His Gly
                 190                 195                 200

AAC ACG GCC AGG CTT GCG AAG GAG CAT GGT GAC ATA AAA TTG GCA   810
Asn Thr Ala Arg Leu Ala Lys Glu His Gly Asp Ile Lys Leu Ala
                 205                 210                 215
```

```
CAG ATC TGC GGC ATG ATT GCC TCA GAT GAG AAG CGC CAC GAG ACT   855
Gln Ile Cys Gly Met Ile Ala Ser Asp Glu Lys Arg His Glu Thr
                220                 225                 230

GCA TAC ACA AAG ATA GTG GAA AAG CTG TTT GAG GTT GAT CCT GAT   900
Ala Tyr Thr Lys Ile Val Glu Lys Leu Phe Glu Val Asp Pro Asp
                235                 240                 245

GGT ACA GTT ATG GCA TTT GCC GAC ATG ATG AGG AAG AAG ATT GCT   945
Gly Thr Val Met Ala Phe Ala Asp Met Met Arg Lys Lys Ile Ala
                250                 255                 260

ATG CCA GCA CAC CTT ATG TAT GAC GGC CGC GAC GAC AAC CTG TTT   990
Met Pro Ala His Leu Met Tyr Asp Gly Arg Asp Asp Asn Leu Phe
                265                 270                 275

GAT AAC TAC TCT GCC GTC GCG CAG CGC ATT GGG GTC TAC ACT GCA  1035
Asp Asn Tyr Ser Ala Val Ala Gln Arg Ile Gly Val Tyr Thr Ala
                280                 285                 290

AAG GAC TAT GCT GAC ATA CTC GAA TTT CTG GTG GGG AGG TGG AAG  1080
Lys Asp Tyr Ala Asp Ile Leu Glu Phe Leu Val Gly Arg Trp Lys
                295                 300                 305

GTG GAG CAG CTA ACC GGA CTT TCA GGT GAG GGA AGA AAG GCT CAG  1125
Val Glu Gln Leu Thr Gly Leu Ser Gly Glu Gly Arg Lys Ala Gln
                310                 315                 320

GAA TAC GTT TGT GGG CTG CCA CCA AGA ATC AGA AGG TTG GAG GAG  1170
Glu Tyr Val Cys Gly Leu Pro Pro Arg Ile Arg Arg Leu Glu Glu
                325                 330                 335

AGA GCT CAA GCA AGA GGC AAG GAG TCG TCA ACA CTT AAA TTC AGT  1215
Arg Ala Gln Ala Arg Gly Lys Glu Ser Ser Thr Leu Lys Phe Ser
                340                 345                 350

TGG ATT CAT GAC AGG GAA GTA CTA CTC TAAATGCT TGCACCAAGG      1260
Trp Ile His Asp Arg Glu Val Leu Leu
                355             359

GAGGAGCATG GTGAATCTTC CAGCAATACC ATTCTGAGAA ATGTTGAATA       1310

GTTGAAAATT CAGTTTGTCA TTTTTATCTT TTTTTTCTCC TGTTTTTTGG       1360

TCTTATGTTA TATGCCACTG TAAGGTGAAA CAGTTGTTCT TGCATGGTTC       1410

GCAAGTTAAG CAGTTAGGGG CAGCTGTAGT ATTAGAAATG CTATTTTTTG       1460

TTTCCCTTTT CTGTGGTAGT GATGTCTGTG GAAGTATAAG TAAACGTTTT       1510

TTTTTTCTC TGGCAATTTTG ATGATAAAGA AAATTTAGTT CTAAAAACCG       1560

TCGCACCTTC CCTGAGGCTT CTCTTGTCTG TCGCGAGTGA CCATGGTGAG       1610

GGTTAGTGTG CTGAACGATG CTCTGAAGAG CATGTACAAT GCTGAGAAAA       1660

GGGGAAAGCG CCAAGTCATG ATTCGGCCAT CCTCCAAAGT CATTATCAAA       1710
```

TTCCTTTTGG TGATGCAGAA GCACGGATAC ATTGGAGAGT TTGAGTATGT 1760

TGATGACCAC AGGGCTGGTA AAATCGTGGT TGAATTGAAC GGTAGACTGA 1810

ACAAGTGTGG GGTTATTAGT CCCCGTTTTG ATGTCGGCGT CAAAGAGATT 1860

GAAGGTTGGA CTGCTAGGCT TCTCCCCTCA AGACAGTTTG GGTATATTGT 1910

ATTGACTACC TCTGCCGGCA TCATGGATCA CGAAGAAGCT AGGAGAAAAA 1960

ATGTTGGTGG TAAGGTACTG GGTTTCTTCT ACTAGAGTTT AATTTCGATT 2010

AAGAGGATGT CAGGAATTTC AATTGAGATT CATGGATTGT AATGGAGGAT 2060

ATGCTAGGCC CCTAGTAATA TCAAGCATAG CAGGAGCTGT TTTGTGATGT 2110

TCCTTATTTT GTTTGCAAAA CCAAGTTGGT AACTATAACT TTTATTTTCT 2160

TTTATCATTA TTTTTCTTTA TACCAAAATG TACTGGCCAA GTTGTTTTAA 2210

ACAGTGAGAA CTTTGATTAG AAAAAAAAAA AAA 2243

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH:216 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: No

(iv) ANTISENSE: No

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Glycine max
 (B) STRAIN: Cultivar Wye
 (D) DEVELOPMENTAL STAGE: Developing seeds

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: cDNA to mRNA
(B) CLONE: pDS4a

(ix) FEATURE:
(A) NAME/KEY:3' non-coding sequence
(B) LOCATION: nucleotides 1 through 216
(C) IDENTIFICATION METHOD: Homology of clones pDS4a and pDS1 and similarity of sequence in SEQ ID NO:1 to 3' non-coding sequence in SEQ ID NO:1

(x) PUBLICATION INFORMATION: Sequence not published.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
GAAATGTTGA ATAGTTGAAA ATTCAGTTTG TCATTTTTAT CTTTTATTTT  50
TTCTCCTTTT TTGGTCTTTG TTATATGTCA CTGTAAGGTG AAGCAGTTGT 100
TCTTGCATGG TTCGCAAGTT AAGCAGTTAG GGGCAGCTGT AGTATTAGAA 150
ATGGTATTTT TTTTTTTGTT TTCGCTTTTC TCTGTGGTAG TGATGTCTGT 200
CGAAGTATAA GTAAAC                                      216
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: No

(v) FRAGMENT TYPE: N-terminal fragment

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Glycine max
(B) STRAIN: Cultivar Wye
(C) DEVELOPMENTAL STAGE: Developing seeds

(ix) FEATURE:
(A) NAME/KEY: N-terminal sequence
(B) LOCATION: 1 through 16 amino acid residues
(C) IDENTIFICATION METHOD: N-terminal amino acid sequencing

(x) PUBLICATION INFORMATION: Sequence not published

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Arg Ser Gly Ser Lys Glu Val Glu Asn Ile Lys Lys Pro Phe Thr Pro
1               5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid: mixture of oligonucleotides

(iii) HYPOTHETICAL: Yes

(ix) FEATURE:

(A) NAME/KEY: Coding sequence

(B) LOCATION: 1 through 36 bases

(x) PUBLICATION INFORMATION : Sequence not published

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
AAR GAR GTN GAR AAY ATH AAR AAR CCN TTY ACN CCN   36
Lys Glu Val Glu Asn Ile Lys Lys Pro Phe Thr Pro
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:35 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid: mixture of synthetic oligonucleotides

(ix) FEATURE:

(C) OTHER INFORMATION: N at positions 3,6,9, and 27 is deoxyinosine.

(x) PUBLICATION INFORMATION: Sequence not published

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGNGTNAANG GCTTCTTRAT RTTYTCNACN TCCTT 35

CLAIMS

What is claimed is:

1. A nucleic acid fragment comprising a nucleotide sequence encoding the soybean seed stearoyl-ACP desaturase cDNA corresponding to the nucleotides 1 to 2243 in SEQ ID NO:1, or any nucleic acid fragment substantially homologous therewith.

2. A nucleic acid fragment of Claim 1 wherein said nucleotide sequence encodes the soybean seed stearoyl-ACP desaturase precursor corresponding to nucleotides 70-1245 in SEQ ID NO:1, or any nucleic acid fragment substantially homologous therewith.

3. A nucleic acid fragment of Claim 2, wherein the said nucleotide sequence encodes the mature soybean seed stearoyl-ACP desaturase enzyme, corresponding to nucleotides 166 to 1245 in SEQ ID NO:1.

4. A chimeric gene capable of transforming a soybean plant cell comprising a nucleic acid fragment of Claim 1 operably linked to suitable regulatory sequences producing antisense inhibition of soybean seed stearoyl-ACP desaturase in the seed.

5. A chimeric gene capable of transforming a plant cell of an oil-producing species comprising a nucleic acid fragment of Claim 2 operably linked to suitable regulatory sequences resulting in overexpression of said soybean seed stearoyl-ACP desaturase in the plastid of said plant cell.

6. A chimeric gene capable of transforming a plant cell of an oil-producing species comprising a nucleic acid fragment of Claim 3 operably linked to suitable regulatory sequences resulting in the expression of said mature soybean seed stearoyl-ACP desaturase enzyme in the cytoplasm of said plant cell.

7. A method of producing soybean seed oil containing higher-than-normal levels of stearic acid comprising:

(a) transforming a soybean plant cell with a chimeric gene of Claim 4,

(b) growing fertile soybean plants from said transformed soybean plant cells,

(c) screening progeny seeds from said fertile soybean plants for the desired levels of stearic acid, and

(d) crushing said progeny seed to obtain said soybean oil containing higher-than-normal levels of stearic acid.

8. A method of producing oils from plant seed containing lower-than-normal levels of stearic acid comprising:

(a) transforming a plant cell of an oil producing species with a chimeric gene of Claims 5 or 6,

(b) growing sexually mature plants from said transformed plant cells of an oil producing species,

(c) screening progeny seeds from said fertile plants for the desired levels of stearic acid, and

(d) crushing said progeny seed to obtain said oil containing lower-than-normal levels of stearic acid.

9. A method of Claim 8 wherein said plant cell of an oil producing species is selected from the group

consisting of soybean, rapeseed, sunflower, cotton, cocoa, peanut, safflower, and corn.

10. A method of Claim 7 wherein said step of transforming is accomplished by a process selected from the group consisting of Agrobacterium infection, electroporation, and high-velocity ballistic bombardment.

11. A method of Claim 8 wherein said step of transforming is accomplished by a process selected from the group consisting of Agrobacterium infection, electroporation, and high-velocity ballistic bombardment.

12. A method of producing mature soybean seed stearoyl-ACP desaturase enzyme in microorganisms comprising:

(a) transforming a microorganism with a chimeric gene of Claim 6,

(b) growing said transformed microorganism to produce quantities of said mature soybean seed stearoyl-ACP desaturase enzyme, and

(c) isolating and purifying said mature soybean seed stearoyl-ACP desaturase enzyme.

13. A method of RFLP breeding of altered levels of stearic acid trait in soybean seed oil comprising:

(a) making a cross between two soybean varieties differing in the trait,

(b) making a Southern blot of restriction enzyme digested genomic DNA isolated from several progeny plants resulting from the cross; and

(c) hybridizing the Southern blot with the radiolabelled nucleic acid fragment of Claim 1.

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 91/03288

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/53 ; C11B1/04 ; A01H1/04 ; C12Q1/68

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	C12N

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category[°]	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
Y	PLANT PHYSIOLOGY vol. 90, 1989, ROCKVILLE USA pages 760 - 764; T. CHEESBROUGH: 'Changes in the enzymes for fatty acid synthesis and desaturation during ac climation of dev.soybean seeds to altered growth temp. ' see the whole document --- -/--	1-13

° Special categories of cited documents : [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
09 SEPTEMBER 1991	07. 10. 91
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	VAN DER SCHAAL C.A.

4

Form PCT/ISA/210 (second sheet) (January 1985)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
Y	CHEMICAL ABSTRACTS, vol. 85, no. 17, October 25, 1976, Columbus, Ohio, US; abstract no. 119683, page 306 ; see abstract & ARCH. BIOCHEM. BIOPHYS. vol. 176, no. 1, 1976, pages 63 - 70; P.K.STUMPF & R. PORRA: 'Lipid biosynthesis in developing and germinating soybean cotyledons. The formation of oleate by a soluble stearoyl ACP desaturase '	1-13
Y	JOURNAL OF BIOLOGICAL CHEMISTRY. vol. 257, no. 20, October 25, 1982, BALTIMORE US T.A.MCKEON & P.K.STUMPF: 'Purification and characterization of the stearoyl-ACP desaturase and the acyl ACP thioesterase from maturing seeds of Safflower , ' cited in the application pages 12141-12147 see abstract	1-13
Y	TRENDS IN BIOTECHNOLOGY. vol. 5, 1987, CAMBRIDGE GB pages 40 - 46; V. KNAUF: 'The application of genetic engineering to oilseed crops ' see the whole document	1-12
Y	EP,A,255 377 (CALGENE, INC.) February 3, 1988 see abstract see page 26, line 28 - line 37; claims	1-13
Y	NL,A,8 800 794 (WESSANEN NEDELAND B.V.) November 1, 1988 see claims	1-12
Y	BIOTECHNOLOGY vol. 7, March 1989, NEW YORK US pages 257 - 264; S.D.TANKSLEY ET AL: 'RFLP mapping in plant breeding: new tools for an old science ' cited in the application see the whole document	13

-/--

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
Y	GENE. vol. 72, 1988, AMSTERDAM NL pages 45 - 50; A. VAN DER KROL: 'Antisense genes in plants: an overview ' cited in the application see the whole document ---	7
A	BIOTECHNOLOGY vol. 6, August 1988, NEW YORK US pages 915 - 922; M.A.W. HINCHEE ET AL.: 'Production of transgenic soybean plants using agrobacterium-mediated DNA transfer ' cited in the application see the whole document ---	10,11
P,Y	WO,A,9 012 084 (DNA PLANT TECHNOLOGY CORPORATION) October 18, 1990 see page 51; example 8 see claims ---	1-13

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO.

US 9103288
SA 48035

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

09/09/91

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-255377	03-02-88	AU-B- 609724	09-05-91
		AU-A- 7630387	04-02-88
		AU-B- 612326	11-07-91
		AU-A- 7630287	04-02-88
		EP-A- 0255378	03-02-88
		JP-A- 63119680	24-05-88
		JP-A- 63112987	18-05-88
NL-A-8800794	01-11-88	NL-A- 8700783	01-11-88
		AU-A- 1415388	06-10-88
WO-A-9012084	18-10-90	AU-A- 5412390	05-11-90
		WO-A- 9011682	18-10-90